

PARAMETER AND VARIANCE ESTIMATION IN GEOTECHNICAL BACKANALYSIS USING PRIOR INFORMATION

A. LEDESMA, A. GENS AND E. E. ALONSO

*Department of Geotechnical Engineering, ETS Ingenieros de Caminos, Technical University of Catalonia (UPC),
c/ Gran Capitán s/n, D-2; 08034 Barcelona, Spain*

SUMMARY

A probabilistic framework to perform inverse analysis of geotechnical problems is presented. The formulation allows the incorporation of existing prior information on the parameters in a consistent way. The method is based on the maximum likelihood approach that allows a straightforward introduction of the error structure of field measurements and prior information. The difficulty of ascribing definite values to the uncertainties associated with the various types of observations is overcome by including the corresponding variances in the set of parameters to be identified. The inverse analysis results in a minimization problem that is solved by coupling the optimization technique to the finite element method. Two examples are presented to illustrate the performance of the method. The first one corresponds to a synthetic case simulating the excavation of a tunnel. Young's modulus, K_0 value and measurements variances are identified. The second case concerns the excavation of a large underground cavern in which again Young's modulus and K_0 are identified. It is shown that introduction of prior information permits the estimation of parameters more consistent with all available informations that include not only monitored displacements but also results from *in situ* tests carried out during the site investigation stage.

KEY WORDS: inverse analysis; parameter identification; underground excavation; optimization technique; finite element method

INTRODUCTION

In Geotechnical Engineering, inverse analysis procedures have been commonly used to determine soil or rock properties from field instrumentation. However, inverse analysis is usually performed in an *ad hoc* manner leading to estimated parameters the reliability of which is generally unknown. Nevertheless, recent developments of numerical analysis procedures coupled with optimization techniques have permitted the formulation of methods for performing geotechnical backanalysis in a more systematic manner.^{1–3} These methods can be considered as included in the general theory of inverse analysis or system identification.^{4–6}

These procedures are very useful when the inverse analysis solution can be employed to improve the geotechnical design of our problem. For instance, in underground excavations, information provided by measurements performed during the first stages of the construction can be used to update the values of the parameters of rock or soil and therefore to assist in design modifications. Also, in time dependent problems,^{7,8} the information available at the beginning of construction can be used to perform a partial inverse analysis and update the values of the parameters needed to make new predictions. The value of systematic inverse analysis can be

significantly enhanced if the information introduced in the analysis is not limited to field measurements performed during construction. In fact, in any project there will exist a significant amount of information on the parameters obtained during the site investigation stage. Incorporation of this 'prior' information into the inverse analysis procedure will result in a better parameter estimation and in a more rational use of all the data available.

Introduction of prior information into the process of inverse analysis is more conveniently performed using a probabilistic framework. In fact, such a formulation carries important additional advantages. For instance, observation errors, reliability of the estimated parameters and of the previous information can be considered in a consistent way.

A few methods available in the literature are suitable for geotechnical inverse analysis:

- (a) *Least-squares method*. The best parameters are those that minimize a function which depends on the square differences between measured and computed variables.
- (b) *Maximum likelihood approach*. The measurements \mathbf{x}^* are considered random quantities and the probability density function of \mathbf{x}^* given a set of parameters \mathbf{p} is maximized.¹¹ This procedure estimates those parameters that maximize the probability of observing the measured data (the 'most likely' parameters).
- (c) *Bayesian approach*. The parameters, \mathbf{p} , are considered random variables and the estimation is performed by maximizing the probability density function of \mathbf{p} given a set of measured values \mathbf{x}^* .⁸⁻¹⁰ This is also called the maximum *a posteriori* approach, because it estimates the parameters that are the most probable from the existing data.
- (d) *Kalman filter approach*. Filters are data processing schemes that separate desired signals from unwanted disturbances (noise).⁴ A general state equation relating measurements, parameters and noise (characterized in statistical terms) is adopted and an optimal estimation of the parameters is proposed in a recursive form.

Although there are conceptual differences between these approaches, the final mathematical formulations are similar in most cases. For instance, the least-squares method can be considered as a particular case of the maximum likelihood approach when error measurements are assumed to be independent and Gaussian with the same variance, and there is no prior information on the parameters. Also, under some conditions, an extension of the Kalman filter method has been proved to be equivalent to the Bayes approach.² Therefore, selecting between approaches (b), (c) or (d) is more a matter of viewpoint, as in many cases final expressions are the same.

In this paper, a general formulation to perform geotechnical inverse analysis based on the maximum likelihood approach is presented. The formulation allows the introduction of the error structure of the observations and gives a minimum bound of the variances of the parameters identified. The formulation results in a minimization problem of an 'objective function', J , which can be solved by means of any suitable optimization technique. To assure a wide applicability of the formulation in practical geotechnical problems, the finite element method has been chosen as the basic procedure to perform direct analyses, and therefore it has been coupled to the optimization scheme.

One of the drawbacks of the probabilistic approaches is that usually some of the probabilistic variables involved in the procedure are not easily determined. For instance, it is difficult to know the measurements variance relative to the error associated with prior information, as they usually are obtained from quite different sources. For this reason, a general procedure in which variances are introduced as additional parameters to be identified is presented. In this way, material parameters and relative variances are jointly estimated in a consistent framework.

The performance of the variance and parameter identification procedure is demonstrated using a synthetic example simulating a tunnel excavation. The effect of introducing 'prior information'

in the analysis is illustrated in an application to a real case involving the staged excavation of a cavern in which both material parameters and relative variances between field measurements and prior information are identified.

MAXIMUM LIKELIHOOD FORMULATION

Basic formulation

It is assumed that a relation between state variables, \mathbf{x} , and parameters, \mathbf{p} , has been defined by means of a model, \mathbf{M} (generally non-linear) which is considered fixed: $\mathbf{x} = \mathbf{M}(\mathbf{p})$. The information available includes some measurements, that is a set of measured state variables, \mathbf{x}^* , and some prior information on the parameters to be estimated, \mathbf{p}^* .

Using the maximum likelihood approach, the best estimation of the system parameters is found by maximizing the likelihood of an hypothesis, L . Likelihood is defined as proportional to the joint probability of errors in measuring state variables and in the prior information of the parameters:

$$L = kP(\mathbf{x}, \mathbf{p}) = kP(\mathbf{x})P(\mathbf{p}) \quad (1)$$

where measured state variables and parameters have been considered as independent.

In fact, we assume that the model is correct and differences between measured values, \mathbf{x}^* , and the values computed using the model, \mathbf{x} , are due to the error measurement process. Also, differences between the prior information on the parameters, \mathbf{p}^* , and the parameters to be estimated, \mathbf{p} , are due to the error in that prior information. Therefore, when equation (1) is used, we are maximizing the probability of reproducing the errors we have obtained in the measurement process and in the generation of prior information.

If probability distributions are supposed to be multivariate Gaussian, then

$$P(\mathbf{x}) = \frac{1}{\sqrt{((2\pi)^m |C_x|)}} \exp\left[-\frac{1}{2}(\mathbf{x}^* - \mathbf{x})'(C_x)^{-1}(\mathbf{x}^* - \mathbf{x})\right] \quad (2a)$$

$$P(\mathbf{p}) = \frac{1}{\sqrt{((2\pi)^n |C_p^0|)}} \exp\left[-\frac{1}{2}(\mathbf{p}^* - \mathbf{p})'(C_p^0)^{-1}(\mathbf{p}^* - \mathbf{p})\right] \quad (2b)$$

where $(\mathbf{x}^* - \mathbf{x})$ is the vector of differences between measured and computed values using a fixed model, $(\mathbf{p}^* - \mathbf{p})$ is the vector of differences between prior information and parameters to be estimated, C_x is the measurements covariance matrix, which represents the structure of the error measurements, C_p^0 is the *a priori* parameters covariance matrix, which represents the error structure of the available prior information, m is the number of measurements, n is the number of parameters, $()'$ is used to indicate a transposed matrix, and $||$ is the determinant symbol.

Note that in (1), $L = L(\mathbf{p})$, i.e. likelihood depends only on parameters, because the relationship $\mathbf{x} = \mathbf{M}(\mathbf{p})$ is introduced in the probability density function (2a). Maximizing L is equivalent to minimize the function $S = -2 \ln L(\mathbf{p})$, that is

$$S(\mathbf{p}) = (\mathbf{x}^* - \mathbf{x})'(C_x)^{-1}(\mathbf{x}^* - \mathbf{x}) + (\mathbf{p}^* - \mathbf{p})'(C_p^0)^{-1}(\mathbf{p}^* - \mathbf{p}) \\ + \ln |C_x| + \ln |C_p^0| + m \ln(2\pi) + n \ln(2\pi) - 2 \ln k \quad (3)$$

where the last three terms are constant and can be disregarded in the minimization process.

Expression (3) shows that the function to be minimized (called 'objective function') depends on the error structure of measurements and prior information through the covariance matrices which are usually difficult to define. Generally, the information available will not be sufficient to

specify all the elements of the covariance matrices and some terms will have to be fixed. To do that, it is convenient to separate measurements and prior information in groups with independent covariance matrices. For instance, if m measurements have been obtained from r independent instruments and n parameters can be divided in s groups with individual *a priori* covariance matrices, then the objective function, S , becomes

$$S(\mathbf{p}) = \sum_{i=1}^r (\mathbf{x}^* - \mathbf{x})'(\mathbf{C}_x)_i^{-1}(\mathbf{x}^* - \mathbf{x}) + \sum_{j=1}^s (\mathbf{p}^* - \mathbf{p})'(\mathbf{C}_p^0)_j^{-1}(\mathbf{p}^* - \mathbf{p}) \\ + \sum_{i=1}^r \ln|(\mathbf{C}_x)_i| + \sum_{j=1}^s \ln|(\mathbf{C}_p^0)_j| \quad (4)$$

It should be pointed out that when no prior information is available, only the first and third terms in (4) must be considered. Moreover, if measurements are independent and all of them have the same variance, we obtain $\mathbf{C}_x = \sigma^2 \mathbf{I}$ where \mathbf{I} is the identity matrix. If the value of σ^2 is fixed in the process, only the first term in equation (4) is relevant and a classical 'least-squares' criterion is obtained from that expression.

Variance estimation

It is convenient to express each individual covariance matrix as

$$(\mathbf{C}_x)_i = \sigma_i^2 (\mathbf{E}_x)_i, \quad (\mathbf{C}_p^0)_j = \sigma_j^2 (\mathbf{E}_p^0)_j \quad (5)$$

where σ^2 plays the role of a scale factor which represents the global variance of the data, whereas the (\mathbf{E}_x) and (\mathbf{E}_p^0) matrices represent the error structure associated with that particular type of data.

Generally, the error structure is constant, that is, it depends only on the measurement instrument or on the procedure used to obtain the prior information on the parameters. Covariance matrices corresponding to a variety of geotechnical instruments have been derived elsewhere.¹²

The global variances σ^2 could be determined from the standard error of the measurement device, but in general there are uncontrolled factors that influence that value, for instance, operator skill or equipment conditions. Therefore, those values are difficult to determine in practice, not only for the measurement covariance matrices but usually also for the prior information covariance matrices. Hence, it is convenient to consider those global variances as additional parameters to be identified. This is consistent with the maximum likelihood approach. It was initially proposed by Theil¹³ and has been successfully used in hydrology and in geotechnical problems.^{14,15} Introducing expressions (5) into the objective function (4) leads to

$$S = \sum_{i=1}^r \sigma_i^{-2} (\mathbf{x}^* - \mathbf{x})'(\mathbf{E}_x)_i^{-1}(\mathbf{x}^* - \mathbf{x}) + \sum_{j=1}^s \sigma_j^{-2} (\mathbf{p}^* - \mathbf{p})'(\mathbf{E}_p^0)_j^{-1}(\mathbf{p}^* - \mathbf{p}) \\ + \sum_{i=1}^r m_i \ln \sigma_i^2 + \sum_{j=1}^s n_j \ln \sigma_j^2 + \sum_{i=1}^r \ln|(\mathbf{E}_x)_i| + \sum_{j=1}^s \ln|(\mathbf{E}_p^0)_j| \quad (6)$$

where m_i and n_j are the dimensions of the individual covariance matrices. The last two terms are now constant (assuming fixed the error structure and only variable the global error σ^2) and will not be taken into account in the minimization process.

Any procedure to minimize function (6) can be adopted.¹⁶ Note that the objective function depends now both on the parameters and on the variances. However, the minimization problem

can be simplified uncoupling the estimation of parameters from the estimation of variances. This is convenient from a practical point of view, as conventional formulations are geared to the identification of parameters only. The minimization is then performed in an iterative two-step procedure: particular values of the variances are first selected and the minimization process is carried out to identify only parameters. Variance values are varied according to an independent optimization procedure until the global minimum is obtained.

To show the validity of this approach, let us consider first a set of $\mu_i = \sigma^{*2}/\sigma_i^2$ fixed, where σ^{*2} is any variance taken as reference. This is equivalent to performing a constrained minimization of (6). The extended Lagrangian of function (6) is

$$\sigma^{*2} - \mu_k \sigma_k^2 = 0, \quad k = 1, \dots, (r + s) \quad (7)$$

i.e.

$$\text{Lag} = S + \sum_k^{r+s} \ell_k (\sigma^{*2} - \mu_k \sigma_k^2) \quad (8)$$

where ℓ_k is a Lagrange multiplier. After deriving (8) with respect to the variances, imposing the minimum condition and eliminating ℓ_i , from the equations, the following expressions are obtained:

$$\sigma^{*2} = \frac{J'}{m + n}, \quad \sigma_k^2 = \frac{J'}{\mu_k(m + n)} \quad (9)$$

where

$$J' = \sum_{i=1}^r \mu_i (\mathbf{x}^* - \mathbf{x})^t (\mathbf{E}_x)_i^{-1} (\mathbf{x}^* - \mathbf{x}) + \sum_{j=1}^s \mu_j (\mathbf{p}^* - \mathbf{p})^t (\mathbf{E}_p^0)_j^{-1} (\mathbf{p}^* - \mathbf{p}) \quad (10)$$

Substituting (9) in expression (6) leads to a condition for S in the minimum:

$$S = m + n + \sum_{i=1}^r m_i \ln \left(\frac{J'}{\mu_i(m + n)} \right) + \sum_{j=1}^s n_j \ln \left(\frac{J'}{\mu_j(m + n)} \right) \quad (11)$$

Note that in (11) the dependence of parameters is through J' . Moreover, as expression (11) is a monotonic function of J' , a minimum of J' will minimize S as well.

The general problem was defined as finding a set of parameters, $\hat{\mathbf{p}}$, and variances, $\hat{\boldsymbol{\mu}}$, that minimized (6). Therefore, if $\hat{\boldsymbol{\mu}}$ is found using an independent algorithm, the parameters, $\hat{\mathbf{p}}$, that minimize J' , will also minimize $S(\mathbf{p}, \boldsymbol{\mu})$ for those $\hat{\boldsymbol{\mu}}$ values. This result allows to uncouple the minimization procedure: i.e. by minimizing J' , different sets of $\hat{\mathbf{p}}$ for different values of $\boldsymbol{\mu}$ are found. The values $(\hat{\mathbf{p}}, \hat{\boldsymbol{\mu}})$ that minimize S are obtained via a direct search algorithm.

Parameters reliability

The maximum likelihood formulation provides an statistical framework within which information on the reliability of the parameters estimated can be obtained. For instance, the covariance matrix of the estimated parameters (*a posteriori* covariance matrix) is⁵

$$\mathbf{C}_p = [(\mathbf{C}_p^0)^{-1} + \mathbf{A}^t(\mathbf{C}_x)^{-1}\mathbf{A}]^{-1} \quad (12)$$

where \mathbf{A} is the sensitivity matrix ($m \times n$):

$$\mathbf{A} = \frac{\partial \mathbf{x}}{\partial \mathbf{p}} \quad (13)$$

Expression (12) takes into account both the prior information and the measurement error to estimate the final covariance of the estimated parameters. It should be pointed out that \mathbf{C}_p computed from (12) is a minimum bound of the parameters variance, due to the linearization of the model implicit in (13).⁵

NUMERICAL IMPLEMENTATION

Optimization procedure

The mathematical problem to be solved in order to perform the estimation of parameters is an unconstrained minimization of (6). Following the procedure described above, the simpler expression (10) can be used instead. Using different values of variances, associated values of J' are obtained by minimization of (10). The global minimum of J' will give the parameters and variances finally estimated.

There are a wide range of algorithms available to find the minimum,¹⁶ but Gauss–Newton method is convenient for this kind of objective functions. In case of convergence difficulties, the extension of that method due to Marquardt¹⁷ (Levenberg – Marquardt algorithm) usually gives good results. As the function to minimize is in general non-linear with respect to the parameters, the procedure works iteratively in the parameters' space. Hence, starting from a point in that space, the parameters correction is found by means of⁵

$$\Delta \mathbf{p} = \Delta \mathbf{p}_0 + [\mathbf{A}'(\mathbf{C}_x)^{-1} \mathbf{A} + (\mathbf{C}_p^0)^{-1} + \lambda \mathbf{I}]^{-1} \mathbf{A}'(\mathbf{C}_x)^{-1} [\Delta \mathbf{x} - \mathbf{A} \Delta \mathbf{p}^0] \quad (14)$$

where $\Delta \mathbf{p} = \mathbf{p}^* - \mathbf{p}$, $\Delta \mathbf{x} = \mathbf{x}^* - \mathbf{x}$ and λ is a scalar which controls the convergence process. If $\lambda \rightarrow 0$, the Gauss–Newton method is obtained. As λ increases, $\Delta \mathbf{p}$ tends towards the maximum gradient direction. Near the minimum λ must reach 0 (or a very small value). No general rule exists to decide the value of λ to be used in each iteration. Usually, if J' becomes smaller in an iteration, λ is decreased, reaching zero close to the minimum. However, if J' increases after using equation (14), the value of λ is also increased in order to approach the gradient direction. In the examples presented in this paper, the following criteria has been used: $\lambda = 10$ as initial value, and $\lambda_{n+1} = 10\lambda_n$ if $J'_{n+1} > J'_n$, otherwise $\lambda_{n+1} = \lambda_n/10$, where n is the iteration number.

Finite element implementation

In order to permit the analysis of realistic geotechnical problems, the maximum likelihood formulation has been coupled to a finite element program. In linear static problems the method requires to solve the classical equation

$$\mathbf{K} \mathbf{u} = \mathbf{f} \quad (15)$$

where \mathbf{K} is the global stiffness matrix, \mathbf{u} the vector of nodal displacements, and \mathbf{f} the nodal forces vector. Also,

$$\mathbf{K} = \int_{\Omega} \mathbf{B}' \mathbf{D} \mathbf{B} d\Omega, \quad \mathbf{f} = \int_{\Gamma} \mathbf{N}' \sigma d\Gamma \quad (16)$$

where Ω represents the domain, \mathbf{B} is the geometry matrix relating strains and nodal displacements, \mathbf{D} contains the constitutive law, \mathbf{N} is the shape function matrix, and $\boldsymbol{\sigma}$, for an excavation problem, is the vector of stresses acting on the excavation boundary, Γ .

The optimization algorithm mentioned above requires the computation of the sensitivity matrix, \mathbf{A} , which should be calculated using the Finite Element approximation. Note that measured movements, \mathbf{x}^* , do not necessarily correspond to nodal displacements, \mathbf{u} . Moreover, the measured values sometimes correspond to relative displacements between two points in a direction which does not coincide with the global reference axes (for instance, when classic extensometer devices are used). Therefore, the following general relation between a particular value, x_i , and the related nodal displacements may be written:

$$x_i = \mathbf{L}_i \mathbf{T}_i \mathbf{N}_i \mathbf{u}_i \quad (17)$$

where \mathbf{u}_i is the vector formed by the nodal displacements of the elements involved in that measurement, \mathbf{N}_i is the shape function matrix for the same elements, \mathbf{T}_i is the rotation matrix to transform the components of displacements in the global co-ordinate system to the measurement direction and \mathbf{L}_i is the matrix containing the linear combination of displacements implied by the type of measurement. When relative measurements are performed, $\mathbf{L} = [1, -1]$.

Then, the global vector of measurements will be given as: $\mathbf{x} = \mathbf{L} \mathbf{T} \mathbf{N} \mathbf{u} = \mathbf{R} \mathbf{u}$, and

$$\mathbf{A} = \frac{\partial \mathbf{x}}{\partial \mathbf{p}} = \mathbf{R} \frac{\partial \mathbf{u}}{\partial \mathbf{p}} \quad (18)$$

If a linear elastic constitutive law is used to characterize the material, the sensitivity matrix may be evaluated by derivation of (15) and rearrangement:

$$\mathbf{A} = \frac{\partial \mathbf{x}}{\partial \mathbf{p}} = \mathbf{K}^{-1} \left(\frac{\partial \mathbf{f}}{\partial \mathbf{p}} - \frac{\partial \mathbf{K}}{\partial \mathbf{p}} \mathbf{x} \right) \quad (19)$$

If the parameter involved in the inverse analysis is Young's modulus, then

$$\frac{\partial \mathbf{K}}{\partial E} = \int_{\Omega} \mathbf{B}^t \frac{\partial \mathbf{D}}{\partial E} \mathbf{B} d\Omega \quad \text{and} \quad \frac{\partial \mathbf{f}}{\partial E} = 0 \quad (20)$$

and the derivative of \mathbf{D} with respect to E is easily computed.^{1,2}

In excavation problems, the horizontal stress coefficient, K_0 , is usually a parameter of prime importance and of uncertain value and it is often included in the set of parameters to be estimated. To obtain the required sensitivity matrix, it should be noted that

$$\frac{\partial \mathbf{K}}{\partial K_0} = 0 \quad \text{and} \quad \frac{\partial \mathbf{f}}{\partial K_0} = \int_{\Gamma} \mathbf{N}^t \frac{\partial \boldsymbol{\sigma}}{\partial K_0} dS \quad (21)$$

If the stresses released by the excavation are the initial ones, $\boldsymbol{\sigma}^0$, then

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^0 = (K_0 \sigma_y, \sigma_y, 0), \quad \frac{\partial \mathbf{f}}{\partial K_0} = (\sigma_y, 0, 0) \quad (22)$$

However, it is quite common to have measurements corresponding to different excavation phases. In that case, at an intermediate stage, the stress vector would be $\boldsymbol{\sigma} = \boldsymbol{\sigma}^0 + \Delta \boldsymbol{\sigma}$, where $\Delta \boldsymbol{\sigma}$ represents the increment of stresses during the previous excavation phases. In this case, computation of (21) becomes cumbersome, due to the term $\partial \Delta \boldsymbol{\sigma} / \partial K_0$. To avoid that, an alternative procedure may be used, based on the computation of the displacements of phase j , \mathbf{u}^j as

$(\mathbf{u}^j - \mathbf{u}^{j-1})$, where \mathbf{u}^j indicates the nodal displacements caused by all the excavation phases up to stage j . Then

$$\frac{\partial \mathbf{u}^j}{\partial K_0} = \mathbf{K}_J^{-1} \frac{\partial \mathbf{f}_J}{\partial K_0} - \mathbf{K}_J^{-1} \frac{\partial \mathbf{f}_{J-1}}{\partial K_0} \quad (23)$$

and now $\partial \mathbf{f}_J / \partial K_0$ can be evaluated exactly from (21).

After calculation of the sensitivity matrix, computation of (14) can be performed and, as the minimization problem is non-linear, the optimization process should continue until convergence to the minimum of the objective function.

Non-linear geotechnical constitutive laws introduce additional difficulties in the inverse analysis. In these cases, expressions similar to (19) for computing the sensitivity matrix can be derived only when there is a one-to-one relationship between stresses and strains.^{18,19} The use of elastoplastic models requires calculation of sensitivity matrix by means of numerical approximations²⁰ and it is out of the scope of this paper.

SYNTHETIC EXAMPLE

Variance estimation

A synthetic example simulating the excavation of a tunnel has been used to illustrate the capabilities of the methodology developed. The material is considered to be linear elastic, homogeneous and isotropic. The parameters to be identified in the problem are Young's modulus and the K_0 coefficient. Poisson's ratio is assumed to be 0.49, simulating undrained conditions, and specific weight is 20 kN/m³.

Figure 1 shows the finite element mesh corresponding to this case, considering only half of the geometry due to symmetry. Nodes numbered 1–24 represent the measurement points to be used as input data: horizontal displacements in points 1–15 correspond to measurement values provided by an inclinometer and vertical displacements in points 16–24 are assumed to be obtained from an extensometer. No prior information has been considered in this example.

A basic analysis has been carried out simulating the excavation in one step and using the parameters $E = 10$ MPa and $K_0 = 1$. The values of the displacements at points 1–24 are presented in Table I. In order to check the procedure described to estimate variances, an statistical perturbation is added to the displacements computed in the direct analysis.

As there are two different instruments involved in the measurement process, two variances must be considered: σ_I^2 is the variance corresponding to the horizontal movements, from an inclinometer, and σ_E^2 is the one corresponding to vertical displacements, from an extensometer. For simplicity, it is assumed that the covariance matrices are diagonal, that is

$$(\mathbf{C}_x)_I = \sigma_I^2 \mathbf{I}, \quad (\mathbf{C}_x)_E = \sigma_E^2 \mathbf{I} \quad (24)$$

and therefore all the measurements are considered independent. The purpose of this example is to show the capability of the procedure for estimating the ratio $\mu = \sigma_I^2 / \sigma_E^2$ as well as the parameters used: $E = 10$ MPa and $K_0 = 1$. To do that, three cases have been studied, adding a perturbation of zero mean and different standard deviations to the displacements presented in Table I:

- (1) Perturbation with $\sigma_I = 0.06$ m, $\sigma_E = 0.015$ m, $\mu = 16$.
- (2) Perturbation with $\sigma_I = 0.03$ m, $\sigma_E = 0.03$ m, $\mu = 1$.
- (3) Perturbation with $\sigma_I = 0.015$ m, $\sigma_E = 0.06$ m, $\mu = \frac{1}{16}$.

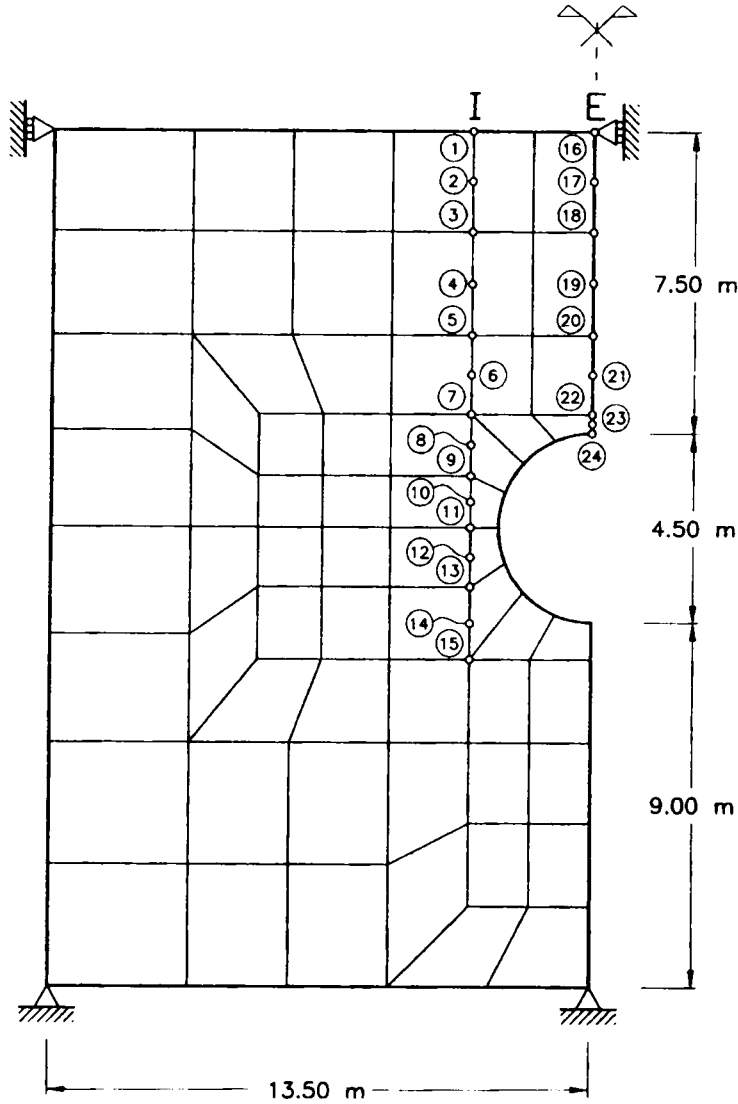


Fig. 1. Finite element mesh and location of the measurement points in the synthetic example

It should be pointed out that the standard deviations used for the perturbation noises are quite important if compared with the displacements indicated in Table I, where the maximum value is about 0.07 m.

The objective function to be minimized is obtained from expression (11). In this case this expression becomes

$$S = 24 + 15 \ln \left(\frac{J'}{24} \right) + 9 \ln \left(\frac{J'}{24\mu} \right) \quad (25)$$

because $m_1 = 15$, $m_2 = 9$, $n = 0$, $r = 2$, $s = 0$, $\mu_1 = 1$, $\mu_2 = \mu = \sigma_1^2 / \sigma_E^2$ as σ_1^2 has been taken as reference variance (σ^{*2}). Note that in (25), $J' = J_I + \mu J_E$, where $J_I = (\mathbf{x}^* - \mathbf{x})_I^T (\mathbf{E}_x)_I^{-1} (\mathbf{x}^* - \mathbf{x})_I$

Table I. Measurements used in the synthetic example

Point	Displacement (cm)
<i>Horizontal movements</i>	
1	0.316
2	0.300
3	0.356
4	0.556
5	0.882
6	1.339
7	2.038
8	2.726
9	3.550
10	4.104
11	4.549
12	4.360
13	3.920
14	3.101
15	2.427
<i>Vertical movements</i>	
16	- 3.598
17	- 3.739
18	- 3.905
19	- 4.119
20	- 4.492
21	- 5.105
22	- 6.328
23	- 6.666
24	- 7.048

using only horizontal displacements, and $J_E = (\mathbf{x}^* - \mathbf{x})_E^T (\mathbf{E}_x)_E^{-1} (\mathbf{x}^* - \mathbf{x})_E$ using vertical displacements only.

For each perturbation case, minimum values of S were determined for a set of different values of μ . Inspection of the results allowed the determination of the global minimum of S . Figure 2 shows that in every case this global minimum of S corresponds to the value of μ of the perturbation applied to the correct measurements. Therefore, the identification procedure has correctly estimated the variance ratio of the data.

The values of the parameters estimated in all analyses are presented in Figure 3. It can be seen that those values are close to the correct ones ($E = 10$ MPa, $K_0 = 1$) and also they are quite similar for any value of μ close to the perturbations applied. Nevertheless, it is interesting to check how good the parameters identified are, using the fact that the real values are known. This can be achieved representing a value of the error $e = [(E - E^*)^2 + (K_0 - K_0^*)^2]^{1/2}$ as a function of μ (Figure 4), where E^* and K_0^* are the actual parameters. In Figure 4, 10 MPa has been taken as unit for Young's modulus so that each parameter has the same weight in calculating e . As Figure 4 shows, the minimum error in the parameters estimated (indicated by a full arrow) is obtained at a value of μ close to the maximum likelihood point (denoted by a dashed arrow). In fact, for case 1, $\mu = 16$, the minimum error parameter point and the global maximum likelihood point coincide. Good estimations for the variance ratio of the measurements and for the parameters involved in the model have been obtained for the joint maximum likelihood identification of parameters and error structure.

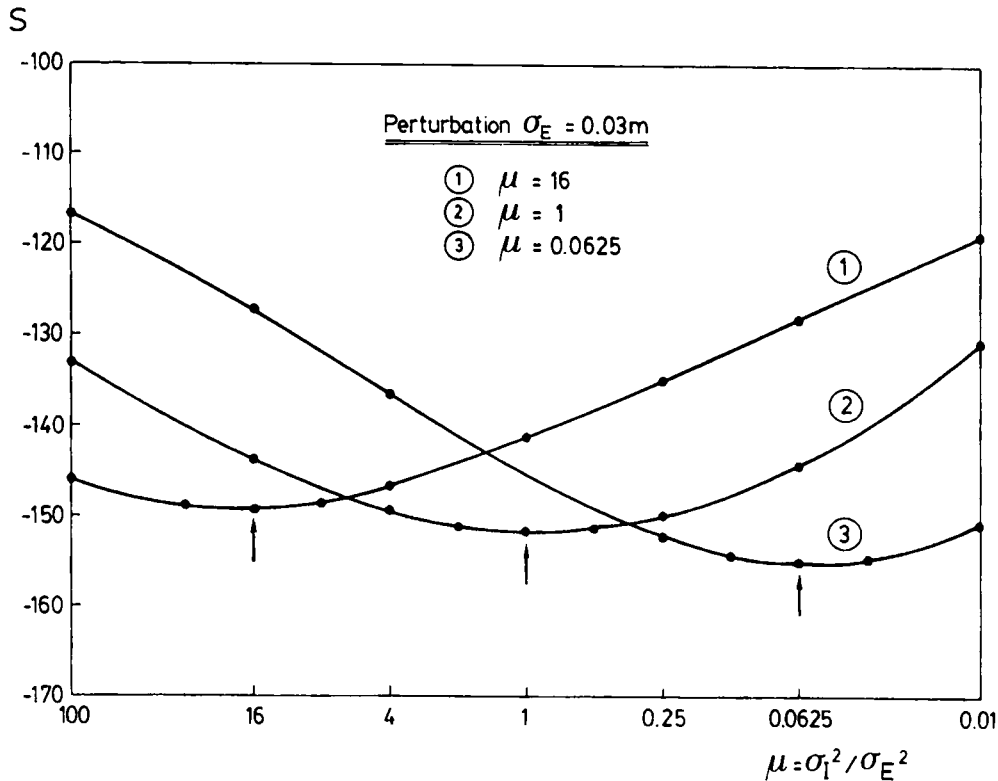


Fig. 2. Variation of objective function value, S , with variance ratio, μ . Arrows indicate the maximum likelihood points which coincide with the minimum of S

Effect of using an incorrect constitutive model

The good performance of the procedure just described may fail if the incorrect constitutive model is selected to represent the mechanic behaviour of soil or rock. To demonstrate this, the same synthetic example has been used. Now the real behaviour of the soil is represented by an anisotropic linear elastic model, but a isotropic law is assumed during the inverse analysis. Undrained conditions, as in the previous example, have been considered.

The anisotropic elastic parameters are:

- (a) Horizontal Young modulus, $E_h = 10$ MPa.
- (b) Vertical Young modulus, $E_v = 5$ MPa.
- (c) Shear modulus (vertical to horizontal) $G_v = 20$ MPa.
- (d) Poisson's ratio (vertical to horizontal) $\nu_{vh} = 0.49$.

The rest of the parameters can be obtained assuming undrained conditions. The K_0 parameter was fixed to 1. Using these parameters, a new set of 24 measurements were obtained through the direct excavation problem. Now the inverse analysis was performed assuming that the material was isotropic but using those displacements computed in anisotropic conditions. Also, all measurements were perturbed with a noise of zero mean and a standard deviation of 0.03 m ($\mu = 1$).

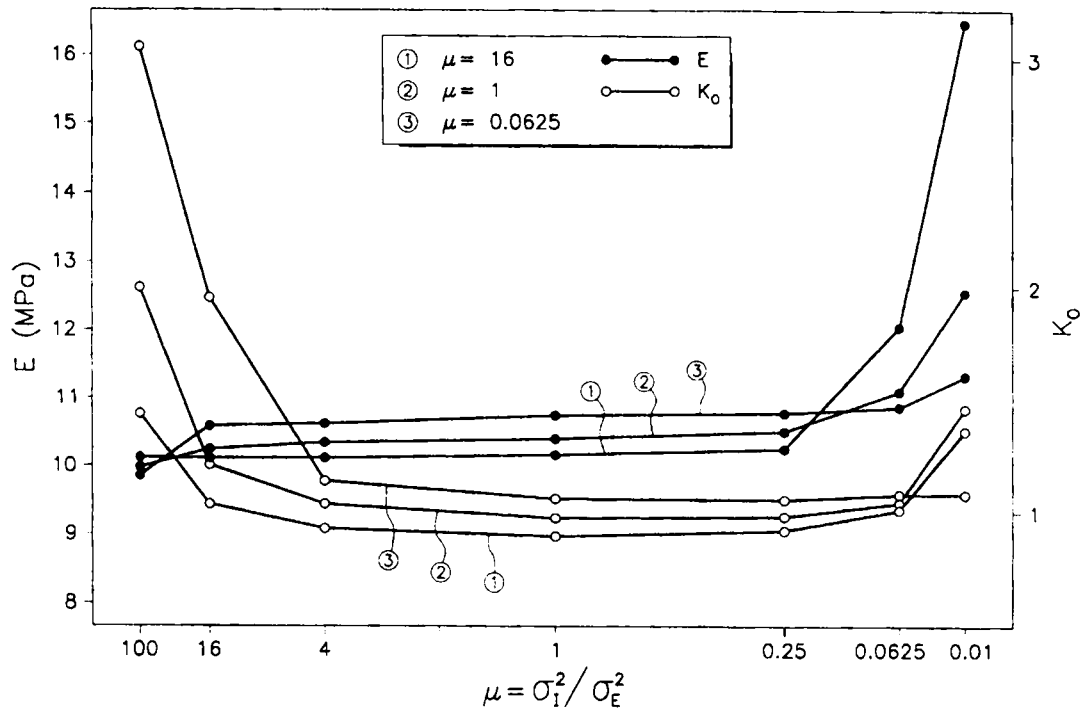


Fig. 3. Parameters estimated for each variance ratio, μ

The identification procedure tried to estimate parameters and variances following the process outlined above. Figure 5 shows the minimum of S obtained for each analysis with fixed μ . It can be observed that the global minimum of S (indicated by a full arrow) is different from the real one ($\mu = 1$). Therefore, the error structure has not been correctly estimated.

The parameters obtained in the inverse analysis (E and K_0) have been depicted with respect to μ in Figure 6. It is not possible to relate the value of the isotropic Young's modulus with the anisotropic moduli. But if the parameter K_0 is considered, neither the maximum likelihood estimate (full arrow), nor the value corresponding to $\mu = 1$ (dashed arrow) are close to the actual value adopted $K_0 = 1$.

This example shows the importance of considering a constitutive model as close to real behaviour as possible, otherwise estimation of variances and parameters may be quite incorrect.

APPLICATION TO A CAVERN EXCAVATION

Description of the problem

As an example of application of the methodology described to a real case, the analysis of the excavation of an underground cavern is presented. The problem involves the construction of the powerhouse cavern of Estanygento-Sallente in the Spanish Pyrennes.^{21,22} The data available from field instrumentation and *in situ* tests have been used to perform inverse analysis in the way described in this paper. The analysis was performed after construction was completed.

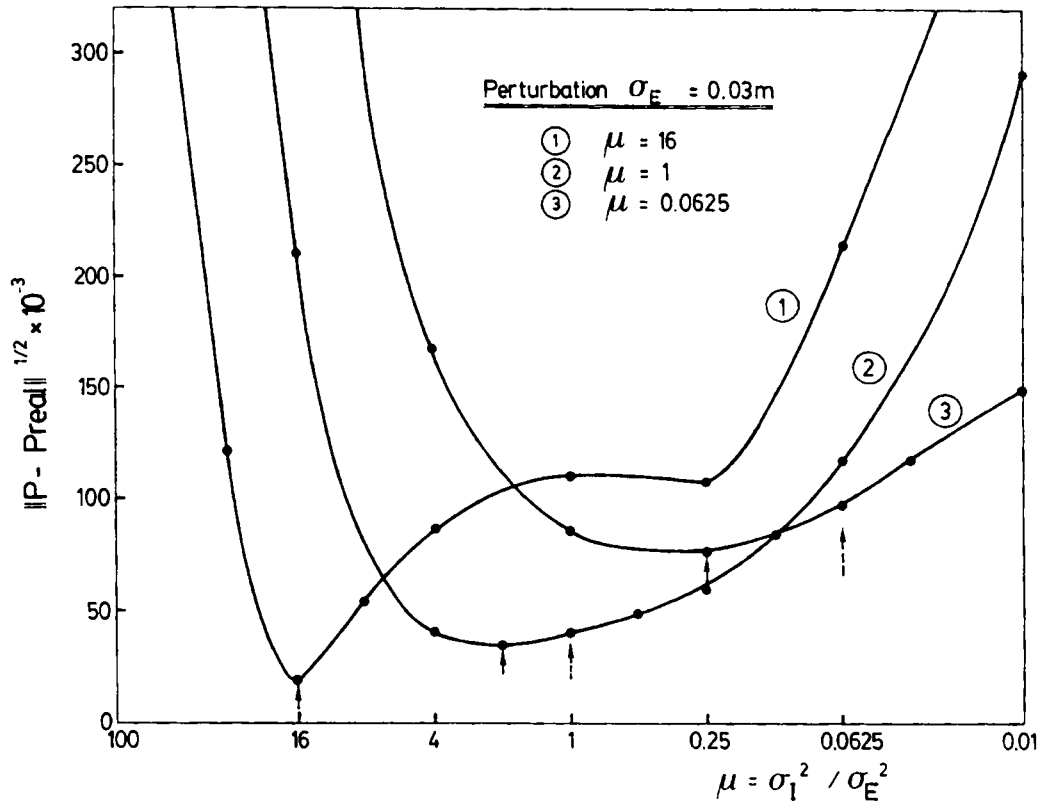


Fig. 4. Estimated parameter error vs. variance ratio μ . Full arrows indicate the maximum likelihood points. Dashed arrows correspond to the minimum parameter error points

The dimensions of the cavern are 37.5 m high, 20 m wide and 89 m long. Plane strain conditions can be assumed if a central section is analysed. Figure 7 shows that section including the location of the measurement points used in this analysis. Some of them (I, J, K, L) are convergence measurements, whereas the rest are displacements obtained from bar extensometers. Therefore, all the measurements available are relative displacements between two points.

The excavation was carried out in several stages, but for the sake of simplicity three phases have been considered in the analysis. Some measurements will be only available for some of the phases. For instance, convergence measuring points K and L (Figure 7) are installed in phase II and therefore measurements are only carried out during excavation of phase III. In contrast, extensometers H – E and M – N provide measurements in all the excavation phases as they were installed from the exploratory tunnel before the excavation of the cavern. Thirty-six measurements have been considered in this analysis: 8 from the first phase, 7 from the second and 21 corresponding to the third phase. Maximum value of a measured relative displacement was 44.9 mm, and the average measurement was 7.7 mm.

In this area the rock is an schist without a definite oriented estratification, due to the high degree of late metamorphism to which it had been subjected. Therefore, the material has been considered homogeneous and isotropic in the analysis, and a linear elastic constitutive law was considered appropriate for the simulation. Hence, the cavern excavation performance depends

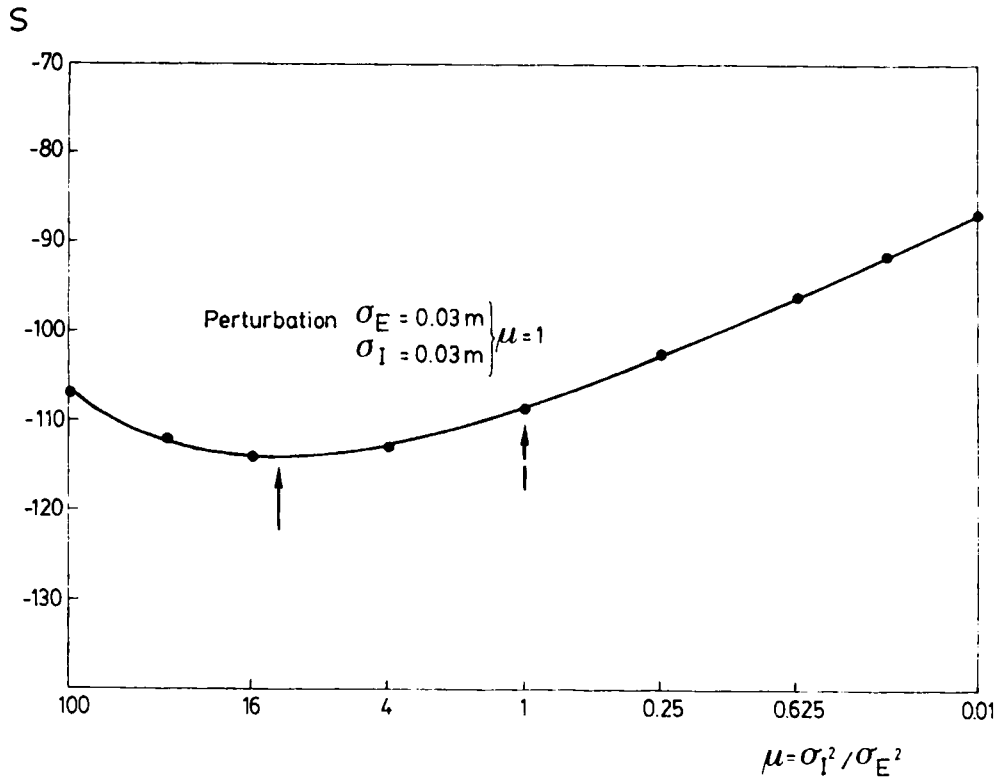


Fig. 5. Variation of S relative to variance ratio μ . Example using an inadequate model. Full arrow indicates the maximum likelihood point, whereas the dashed arrow corresponds to the variance ratio of the real perturbation applied

mainly on Young's modulus of the rock, E , and on the K_0 coefficient. They have been selected as parameters to be identified.

During the site investigation phase, information on E and K_0 was collected that will be included as prior information in the inverse analysis. *In situ* tests such as flat jack tests and dilatometer tests gave an average Young's modulus of $E = 1.5 \times 10^4$ MPa. Those tests confirmed the assumption of isotropy for the rock, but presented some scattering in the value of E . Laboratory tests gave a consistent value of 0.28 for the Poisson's ratio, which is assumed constant and known in the inverse analysis.

In situ stresses were measured in different locations using triaxial inclusion cells and flat jack tests. The results of the tests were consistent with principal stress directions being aligned with the main axis of the cavern. Vertical initial stresses were found to be approximately equal to overburden pressure. However, the values of measured horizontal stresses showed a large scatter and were very dependent on the type of test performed.^{21,22} They became one of the major uncertainties of the project. The value of K_0 obtained ranged from 1 to 3.

Estimation of parameters with no prior information

An analysis for estimating parameters without considering prior information was first performed. As measurements were obtained from bar extensometers, they can be assumed

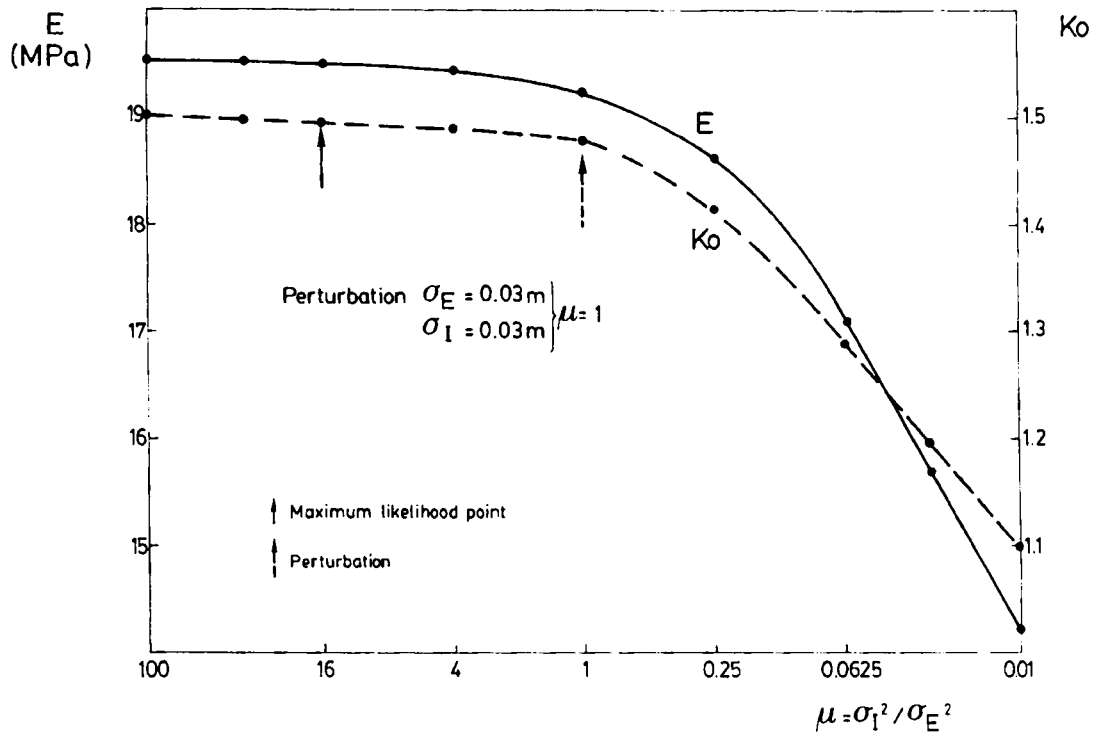


Fig. 6. Parameters estimated for each variance ratio, μ . Example using an inadequate model

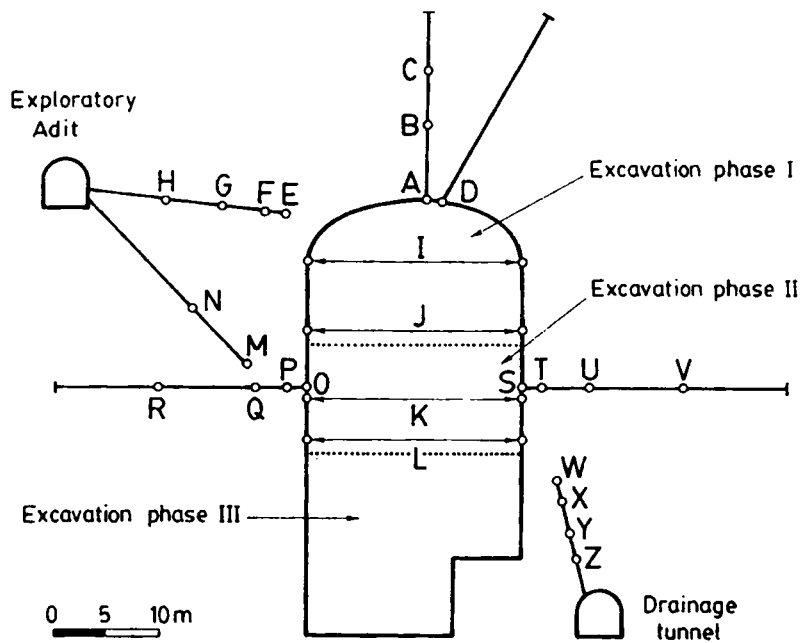


Fig. 7. Underground cavern excavation and field instrumentation

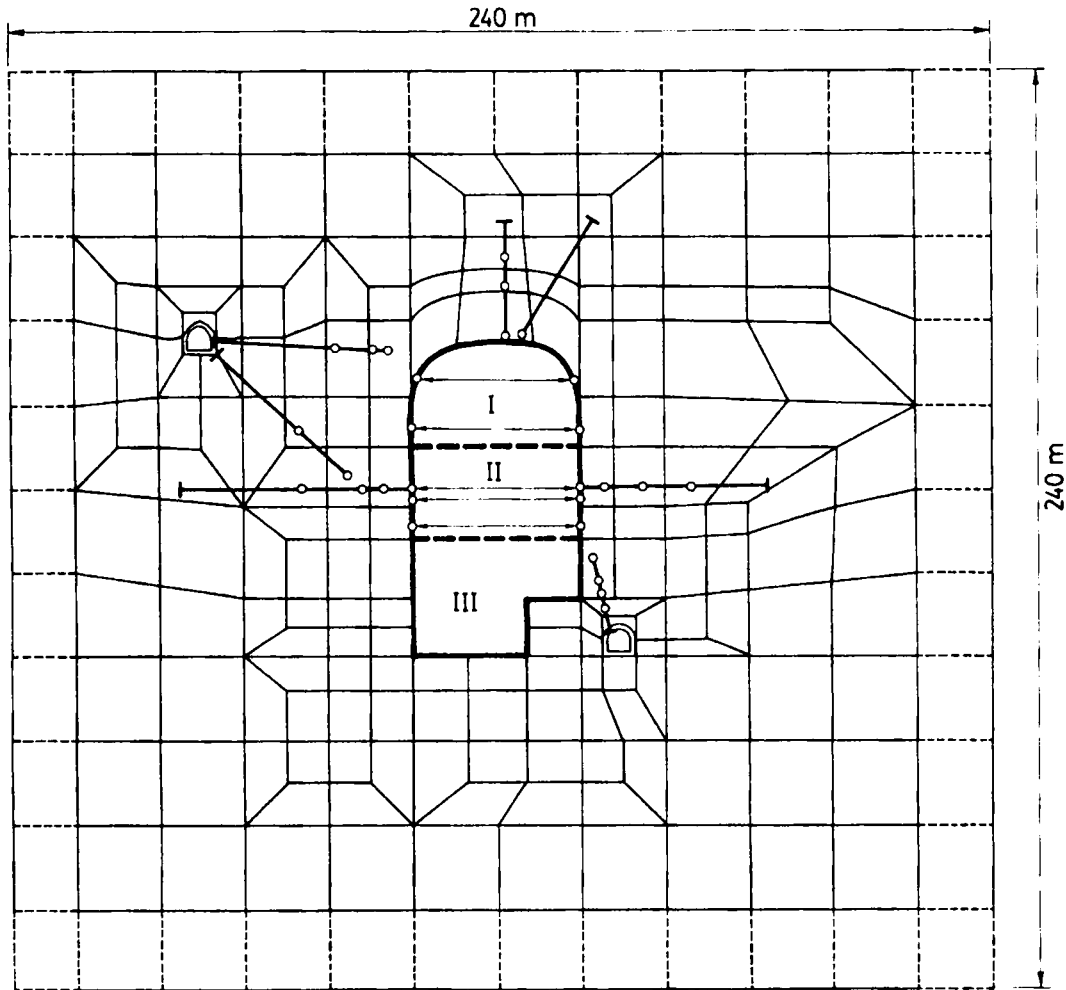


Fig. 8. Finite element mesh used in the analysis of the cavern excavation

independent and therefore, their covariance matrix will be $C_x = \sigma_x^2 I$, where σ_x^2 is the measurement variance. With this assumption the objective function to be minimized is $J = (x^* - x)^T (x^* - x)$ which corresponds to the least-squares criterion. The finite element mesh used is presented in Figure 8. It was checked that changes in dimensions of the mesh produced marginal effects on the relative displacements used as measurements.

The procedure outlined gave the following parameters as estimation from the 36 measurements available: $E = 3900$ MPa, $K_0 = 1.24$. The value of the objective function at the minimum was $J = 2.89 \times 10^{-4} m^2$ which allows the computation of the estimated standard deviation by means of²³

$$\sigma_x = \sqrt{\left(\frac{J}{m - n}\right)} = 2.9 \text{ mm} \quad (26)$$

Figure 9 shows contour levels of the objective function corresponding to this case. Note that the minimum is located in a valley and there are in fact many combinations of E and K_0 that give similar result for J . This can be more clearly seen in Figure 10 where a cross-section of J along line $A-A'$ has been depicted. A comparison between measured displacements and computed values using the parameters identified is presented in Figure 11. The comparison seems reasonable although from Figures 9 and 10 it is clear that other combinations of E and K_0 would also give a comparable agreement.

The identification procedure provides also information on the reliability of the parameters estimated by means of the *a posteriori* parameters covariance matrix. It can be computed from (12), where the term (C_p^0) is not included in this case:

$$C_p = \left[A^t \left(\frac{1}{\sigma_x^2} I \right) A \right]^{-1} \quad (27)$$

Using (26), expression (27) gives the values: $\sigma_E = 1750$ MPa and $\sigma_{K_0} = 0.49$, that are consistent with Figure 9. In fact, the shape of the valley near the minimum in Figure 9 is related to the standard deviations obtained,⁵ although expression (27) is strictly valid only in a local sense, at the minimum, due to the linearization performed in (12) and (13).

Estimation of parameters considering prior information

The parameters estimated in the previous inverse analysis, however, do not appear to be compatible with the prior information obtained from *in situ* tests. In particular, the measured

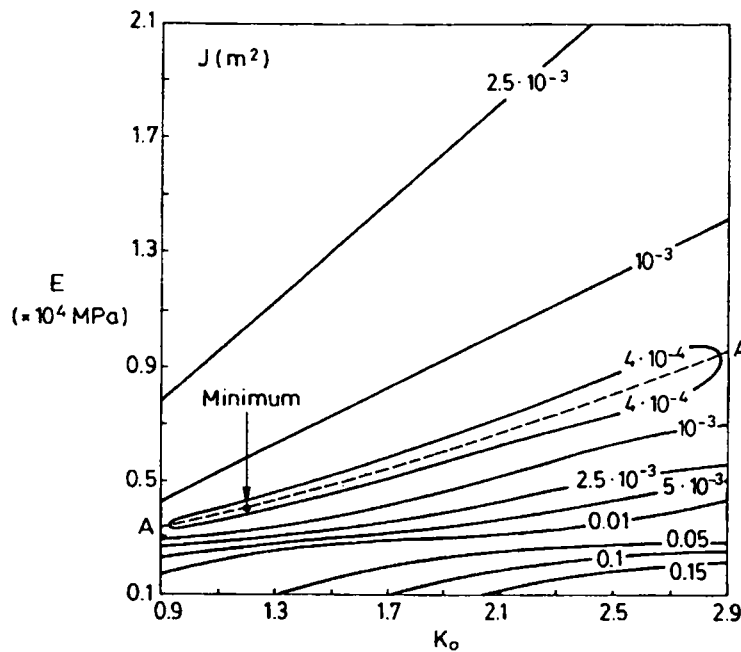


Fig. 9. Objective function contours. Analysis with no prior information

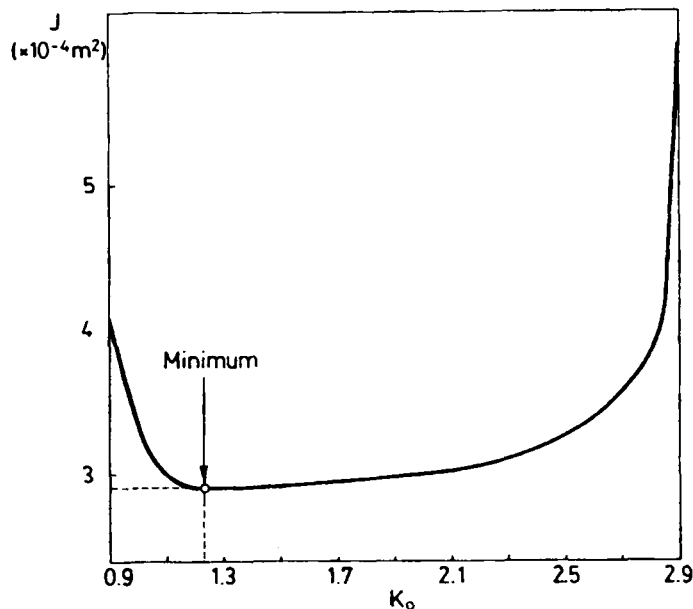


Fig. 10. Section of the objective function along line A – A'

value of E was much greater than the identified one. Therefore, the information provided from those tests should be considered in the inverse analysis as well as measurements, as both are independent estimations of the parameters involved.

When prior information is introduced in the analysis, equation (4) must be used. Prior information on the parameters is defined by means of an average value and a standard deviation of each parameter. The *in situ* tests that were carried out to obtain E and K_0 provided mean values for these parameters: $E = 1.5 \times 10^4$ MPa and $K_0 = 2$. However, standard deviations were difficult to define. This is in general the case, when different procedures in the laboratory or in the field are used to estimate the same parameter. Therefore, in accordance with the methodology described in this paper, variances associated with prior information could also be included in the inverse analysis in the same way as variances associated with measurements have been included in the synthetic examples. As the minimization problem depends in fact on the ratio of variances, the parameters to be identified will be

$$E, \quad K_0, \quad \mu_E = \frac{\sigma_x^2}{\sigma_E^2}, \quad \mu_{K_0} = \frac{\sigma_x^2}{\sigma_{K_0}^2} \quad (28)$$

and the objective function will be similar to expression (6) with $E_x = \mathbf{I}$ and $E_p = \mathbf{I}$, since it is reasonable to assume that all the quantities involved are independent.

Particularizing (11) to this case:

$$S = 36 + 2 + 36 \ln \left(\frac{J'}{38} \right) + \ln \left(\frac{J'}{38 \mu_E} \right) + \ln \left(\frac{J'}{38 \mu_{K_0}} \right) \quad (29)$$

as 2 parameters and 36 measurements are involved in this example.

Figure 12 shows the contours of function (29) near the minimum. It can be seen clearly that the minimum corresponds to $\mu_E = 5 \times 10^{-5} (\text{m}^2/10^8 \text{ MPa}^2)$ and $\mu_{K_0} = 10^{-5} \text{ m}^2$. The parameters

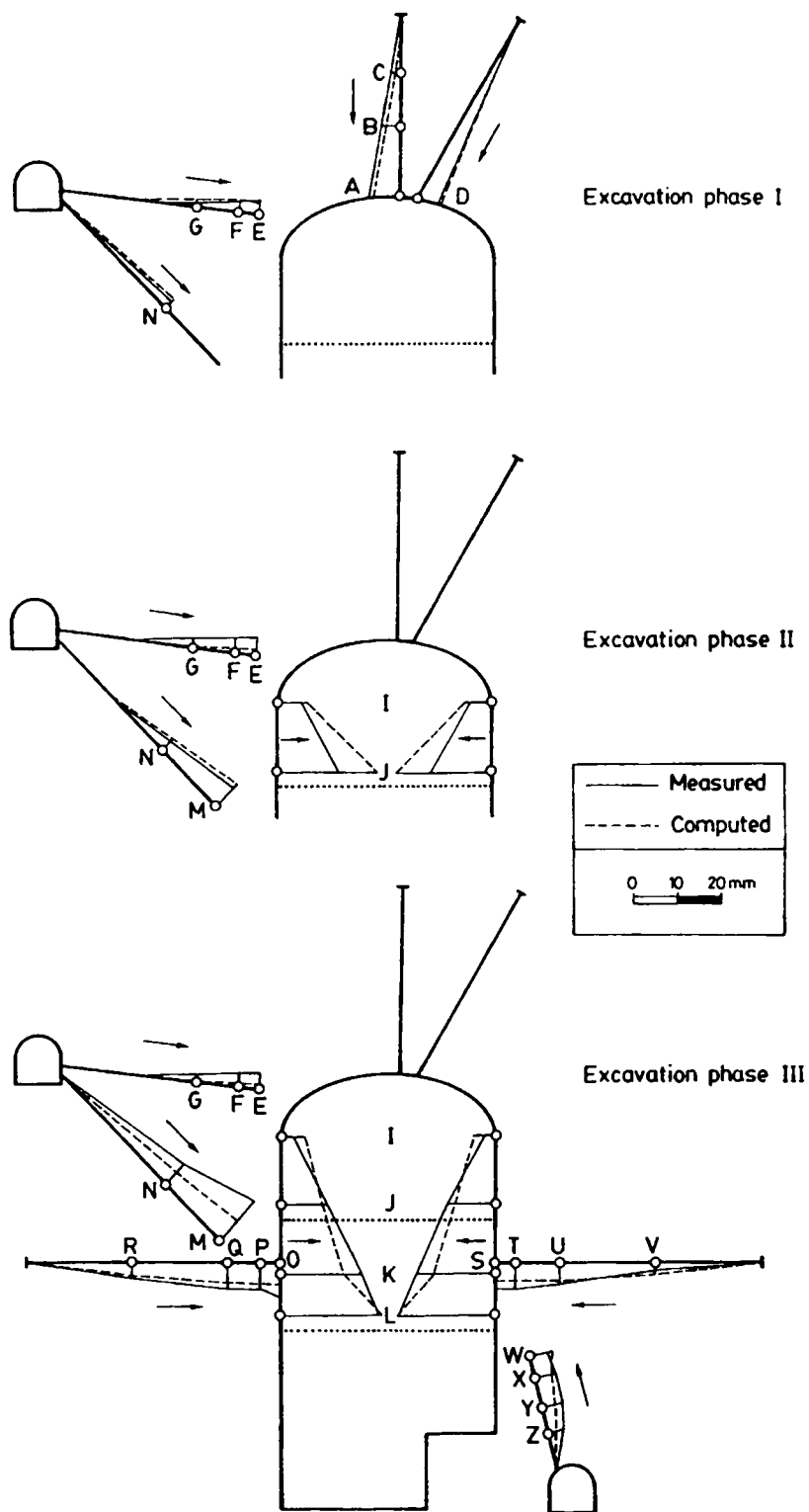


Fig. 11. Comparison between measured and computed displacements. Parameters estimated with no prior information

corresponding to these values are $E = 8200$ MPa and $K_0 = 2.53$, which are much more consistent with the information available. Using expression (26) an estimation of the measurements standard deviation can be obtained: $\sigma_x = 3.1$ mm. The *a posteriori* parameters covariance matrix computed from (12) gives a minimum bound of the reliability of the parameters identified:

$$\mathbf{C}_p = \left[\begin{pmatrix} (\sigma_E^2)^0 & 0 \\ 0 & (\sigma_{K_0}^2)^0 \end{pmatrix}^{-1} + \mathbf{A}' \left(\frac{1}{\sigma_x^2} \mathbf{I} \right) \mathbf{A} \right]^{-1} \quad (30)$$

which results in the values $\sigma_E = 395$ MPa and $\sigma_{K_0} = 0.11$.

Note that the standard deviations obtained in this case are smaller than the previous case, with no prior information. In fact, expression (30) gives always smaller values of \mathbf{C}_p when \mathbf{C}_p^0 is considered, reflecting the reduction of uncertainty brought about by the incorporation of the information given by the site investigation measurements.

Figure 13 presents the comparison between measured and computed displacements using this new set of parameters. The agreement is similar to that obtained in the previous case with no prior information, as the estimated value of σ_x is almost the same. Table II summarizes results obtained for both cases: analyses with and without prior information. Although the error is quite similar, the parameters identified are quite different in each case. The conceptual difference between both solutions is that the latter one incorporates the existing prior information on the parameters with a reliability level which is determined by maximizing the likelihood of the whole system. Therefore, the solution is consistent in a global sense with all the sources of information available: field measurements and prior information from field tests.

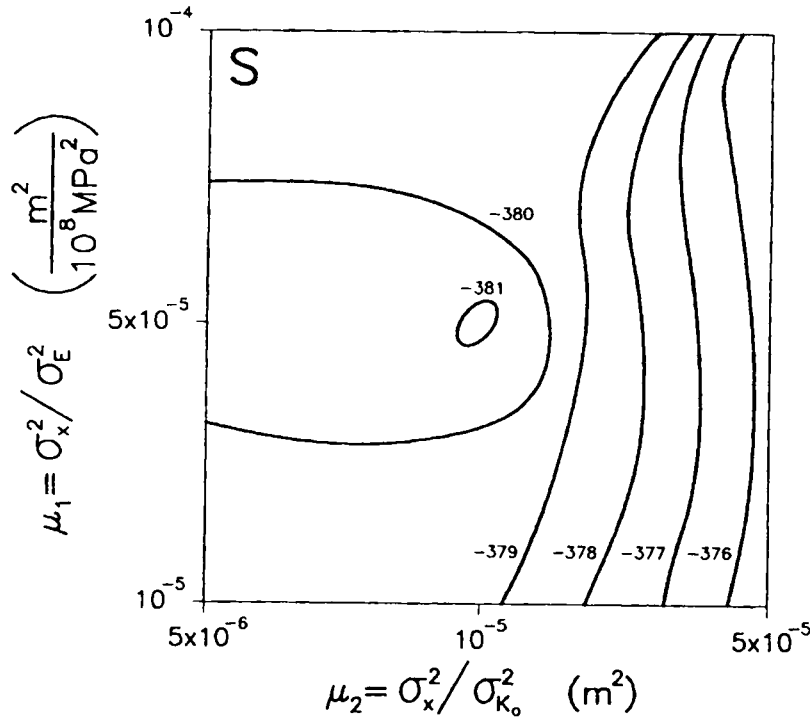


Fig. 12. Contours of objective function defined in terms of variance ratios of E and K_0 .

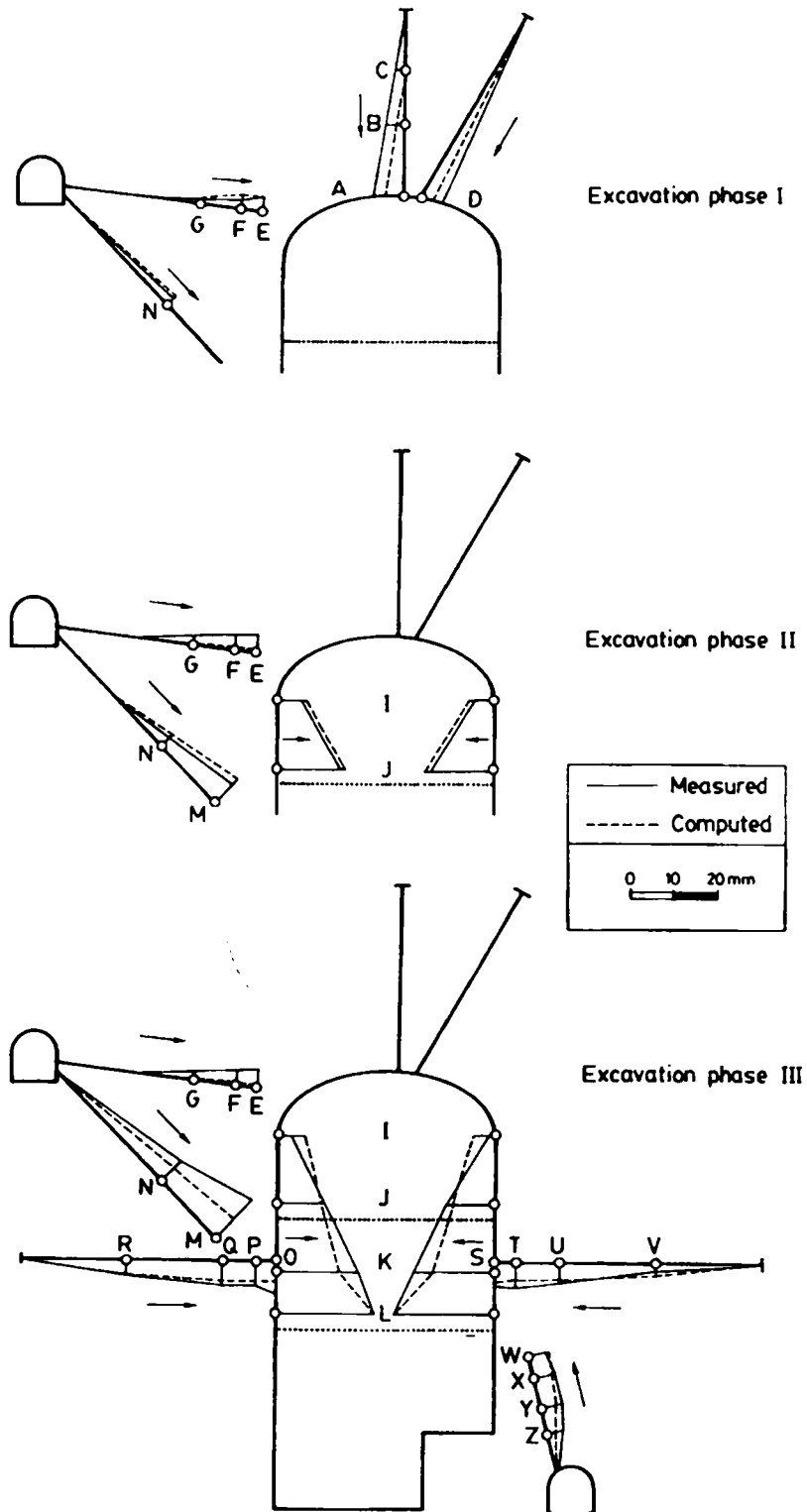


Fig. 13. Comparison between measured and computed displacements. Parameters estimated with prior information

Table II Results obtained in the analysis of a cavern excavation

	E (MPa)	σ_E (MPa)	K_0	σ_{K_0}	σ_x (mm)
No prior information analysis	3900	1750	1.24	0.49	2.9
Prior information analysis	8200	395	2.53	0.11	3.1

CONCLUDING REMARKS

A general formulation to perform geotechnical inverse analysis from field measurements and existing prior information has been presented. The approach is defined using a probabilistic framework and it has the important advantage of allowing the incorporation of different sources of information (field measurements, field and laboratory tests, etc.) in a consistent way.

When this probabilistic framework is used, two main difficulties arise: the relative reliability definition of the contribution of each source to the final estimation and the error structure associated with the input data. To overcome that, the relative weight of measurements and prior information in the final estimation has been considered as an additional parameter to be estimated. Also, the error associated with measurements and to prior information has been included as additional parameters. In this way, the problem is solved considering in a consistent manner all sources of information and their associated errors.

The inverse analysis problem results in a minimization of an objective function which has been performed by means of the Gauss–Newton and Levenberg–Marquardt algorithms. A formulation to use those algorithms together with the Finite Element Method in excavation problems is presented. In order to demonstrate the capabilities of the procedure, a synthetic example relative to the excavation of a tunnel has been presented. The parameters E and K_0 as well as the variance ratio between horizontal and vertical measurements have been identified. The importance of the correct choice of the constitutive model is also highlighted.

Finally, an application to a real problem, involving the excavation of a cavern in rock has been presented. Without prior information the procedure gave parameters from measured displacements that were not consistent with estimations obtained from *in situ* tests. A new identification procedure including prior information has allowed the estimation of E and K_0 for the rock, as well as the relative weight of the prior information with respect to the field measurements. Hence, a new set of parameters has been identified in such a way that all sources of information available (measurements and *in situ* tests) have been considered in the process. Their relative contribution to the final estimation has been determined by the procedure itself in a consistent manner, maximizing the likelihood of the whole system.

ACKNOWLEDGEMENTS

The support given by the Spanish Research Committee 'CICYT' (project PB92-0702) and by FECSA is gratefully acknowledged.

REFERENCES

1. G. Gioda and S. Sakurai, 'Back analysis procedures for the interpretation of field measurements in geomechanics', *Int. j. numer. analyt. methods geomech.* **11**, 555–583, (1987).
2. A. Murakami, 'Studies on the application of Kalman filtering to some geomechanical problems related to safety assesment, *Ph.D. Thesis*, Kyoto University, 1991.
3. Y. Honjo, B. Limanhadi and L. Wen-Tsung, 'Prediction of single pile settlement based on inverse analysis', *Soils Found.*, **33**, 126–144 (1993).
4. P. Eykhoff, *System Identification. Parameter and State Estimation*, Wiley Chichester, 1974.
5. A. Tarantola, *Inverse Problem Theory. Methods for Data Fitting and Model Parameter Estimation*, Elsevier, Amsterdam, 1987.
6. H. D. Bui, *Introduction aux Problèmes Inverses en Mécanique des Matériaux*, Eyrolles, Paris, 1993.
7. K. Arai, H. Ohta and K. Kojima, 'Estimation of soil parameters based on monitored movement of subsoil under consolidation', *Soils Found.* **24**, 95–108 (1984).
8. A. Asaoka and M. Matsuo, 'Bayesian approach to inverse problems in consolidation and its application to settlement prediction', *Proc. 3rd Int. Conf. Num. Meth. Geomechanics*, Aachen, Vol **1**, 1979, pp. 115–123.
9. A. Cividini, G. Maier and A. Nappi, 'Parameter estimation of a static geotechnical model using Bayes' approach', *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.*, **20**, 215–226 (1983).
10. Y. Honjo, L. Wen-Tsung and S. Guha, 'Inverse analysis of an embankment on soft clay by extended bayesian method', *Int. j. numer. analyt. methods geomech.*, **18**, 709–734 (1994).
11. A.W.F. Edwards, *Likelihood*, Cambridge University Press, Cambridge, 1972.
12. A. Ledesma, A. Gens and E. E. Alonso, 'Estimation of parameters in geotechnical backanalysis. I—maximum likelihood approach', *Comput. Geotech.*, **18**, **1**, (1995).
13. H. Theil, 'On the use of incomplete prior information in regression analysis', *Amer. Statist. Soc. J.*, **58**, 401–414 (1963).
14. J. Carrera and S. P. Neuman, 'Estimation of aquifer parameters under transient and steady state conditions. 1, Maximum likelihood method incorporating prior information. 2, Uniqueness, stability and solution algorithms. 3, Application to synthetic and field data.', *Water Resources Res.*, **22**, 199–242 (1986).
15. A. Ledesma, Identificación de parámetros en geotecnia. Aplicación a la excavación de túneles', *Ph.D. Thesis*, Technical University of Catalonia UPC, Barcelona, 1987.
16. R. Fletcher, *Practical Methods of Optimization. 1, Unconstrained Optimization 2, Constrained Optimization*, Wiley, Chichester, 1981.
17. D. W. Marquardt, 'An algorithm for least-squares estimation of nonlinear parameters', *J. Soc. Indust. Appl. Math.* **11**, 431–441 (1963).
18. K. Arai, H. Ohta and K. Kojima, 'Estimation of nonlinear constitutive parameters based on monitored movement of subsoil under consolidation', *Soils Found.*, **27**, 35–49 (1987).
19. A. Ledesma, A. Gens and E. E. Alonso, 'Identification of parameters of nonlinear geotechnical models', *Proc. 7th Int. Conf. on Computer Meth. and Advances in Geomech.*, Cairns Vol. **2**, Balkema, Rotterdam, 1991, pp. 1005–1010.
20. A. Ledesma, J. J. Rosas and A. Gens, 'Identificación de los parámetros de un modelo de estado crítico por retroanálisis', in *Métodos Numéricos en Ingeniería—Spanish Nat. Conf.*, La Coruña, Vol. **1**, SEMNI, 1993, pp. 491–500.
21. W. Wittke and J. L. Soria, 'Exploration, design and excavation of the powerhouse cavern Estanygento-Sallente in Spain', *Proc. 5th Int. Conf. I. S. Rock Mech.*, Melbourne, 1983, pp. D167–178.
22. J. L. Soria and W. Wittke, 'Design and construction of an underground powerhouse in rock subjected to high horizontal stresses', *Proc. 6th Nat. Rock Mech. Symp.*, Achen, 1984, pp. 115–129.
23. R. A. Wiggins, 'The general linear inverse problem: implication of surface waves and free oscillations for earth structure', *Rev. Geophys. Space Phys.*, **10**, 251–285 (1972).